Paleo-islands as refugia and sources of genetic diversity within volcanic archipelagos: The case of the widespread endemic Canarina canariensis (Campanulaceae).

M. MAIRAL*, I. SANMARTÍN†, J. J. ALDASORO2, V. CULSHAW1, I. MANOLOPOULO3 AND M. ALARCÓN†

1 Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain
2 Institut Botànic de Barcelona (IBB-CSIC-ICUB), 08038 Barcelona, Spain
3 University College London, WC1E 6BT London, United Kingdom

Correspondence to be sent to: Real Jardín Botánico (RJB-CSIC), 28014 Madrid, Spain;
Tel: +34914203017; Fax: +34914200157; e-mail: mariomairal@gmail.com
† These senior authors contributed equally

Keywords: ancestral areas, extinction, oceanic islands, paleo-islands, volcanic refugia.

Running head: WITHIN-SPECIES DIVERSIFICATION IN VOLCANIC ISLANDS

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.13282
This article is protected by copyright. All rights reserved.
ABSTRACT

Geographical isolation by oceanic barriers and climatic stability has been postulated as some of the main factors driving diversification within volcanic archipelagos. However, few studies have focused on the effect that catastrophic volcanic events have had on patterns of within-island diversification in geological time. This study employed data from the chloroplast (cpDNA haplotypes) and the nuclear (AFLPs) genomes to examine patterns of genetic variation in Canarina canariensis, an iconic plant species associated with the endemic laurel forest of the Canary Islands. We found a strong geographic population structure, with a first divergence around 0.8 Ma that has Tenerife as its central axis and divides Canarian populations into eastern and western clades. Genetic diversity was greatest in the geologically stable "paleo-islands" of Anaga, Teno and Roque del Conde; these areas were also inferred as the ancestral location of migrant alleles towards other disturbed areas within Tenerife or the nearby islands using a Bayesian approach to phylogeographic clustering. Oceanic barriers, in contrast, appear to have played a lesser role in structuring genetic variation, with intra-island levels of genetic diversity larger than those between-islands. We argue that volcanic eruptions and landslides after the merging of the paleo-islands 3.5 million years ago played key roles in generating genetic boundaries within Tenerife, with the paleo-islands acting as refugia against extinction, and as cradles and sources of genetic diversity to other areas within the archipelago.

INTRODUCTION

Due to their small size, discrete geographic boundaries, substantial environmental heterogeneity and buffered climates, oceanic islands represent excellent natural laboratories to examine the role of ecological adaptations versus geographic isolation as drivers of diversification (Carson & Templeton 1984; Gillespie 2004). Patterns of genetic variation are
expected to be hierarchical in islands, with between-island genetic differentiation stronger than within-islands (Bottin et al. 2005; García-Verdugo et al. 2010), because oceanic barriers are generally more effective than topographic barriers at promoting isolation in insular systems (Gillespie & Clague 2009). However, the birth and development of volcanic islands is usually followed by a large number of destructive events in the form of secondary eruptions, landslides, merging of paleo-islands, etc (Carracedo 1994; Fernández-Palacios et al. 2011). These events promote habitat fragmentation; the subsequent genetic isolation of populations; and they ultimately drive differentiation and speciation within-islands (Carson et al. 1990; Gillespie & Roderick 2002, 2014; Macías-Hernández et al. 2013). Therefore, volcanic archipelagos represent an ideal framework for studying patterns of diversification at different spatial scales (between- and within-islands) and also over different temporal levels (allopatric speciation, extirpations and recolonisations).

The Canary Islands are a volcanic archipelago formed by a chain of seven islands, located 110 km from the north-western coast of Africa. Their geological history and highly endemic biota have interested scientists since the early 19th century (von Humboldt 1814; Lyell 1855). The islands were formed in the last 21 million years (Ma) with an east to west pattern of chronological emergence due to a mantle plume (Carracedo et al. 1998; Zaczek et al. 2015) (Figure 1a). They are separated by deep oceanic trenches and have never been connected to the mainland. The Canary Islands have long been considered refugia for continental lineages that have survived the climatic changes of the Late Cenozoic, but also as cradles of biodiversity where multiple in-situ diversification events have taken place (Juan et al. 2000; Francisco-Ortega et al. 2000). In addition, these islands are regarded as a hotspot for plant diversity (Médail & Quézel 1997): approximately 40% of Canarian vascular plants are endemics (Santos-Guerra 2001) and the flora is generally characterised by high levels of inter-population differentiation in comparison to other archipelagos (Francisco-Ortega et al.)
2000; de Paz & Caujapé-Castells 2013). Recently, they have been proposed as reservoirs and sources of genetic diversity (García-Verdugo et al. 2015; Patiño et al. 2015).

Most studies on the Canarian flora have focused on the pattern of inter-island diversification, particularly at the species level (Francisco-Ortega et al. 2002; Kim et al. 2008; Vitales et al. 2014a, b). Nevertheless, the complex topographies and long histories of avalanches and secondary eruptions of many of these islands are likely to have favoured within-island diversification (Juan et al. 2000; Brown et al. 2006). Tenerife has the most complex history of all the islands. It existed at first as three separate islands, dating back to the Late Miocene: Roque del Conde (11.9-8.9 Ma) in the southwest, Teno (6.2-5.6 Ma) in the northwest, and Anaga (4.9-3.9 Ma) in the northeast. Eruptive central volcanic episodes fused these islands 3.5 Ma and gave rise to the present island of Tenerife (Ancochea et al. 1990; Fig. 1b). The paleo-islands remained thereafter relatively stable, whereas the central part of Tenerife continued to be active until 0.13 Ma (Ancochea et al. 1990, 1999; Cantagrel et al. 1999; Guillou et al. 2004; Carracedo 2014; see Fig. 1b). Interestingly, the three paleo-islands of Tenerife, together with La Gomera – which has also remained geologically stable since the Pliocene – are presently home to the highest phylogenetic diversity and endemic richness of the Canarian Archipelago (Reyes-Betancourt et al. 2008). These areas also harbour the best-preserved laurel forests, considered as an ancient, unique flora restricted to Macaronesia, and they share several endemic and restricted species (see Table S1). Besides their geological (volcanic) stability, the three paleo-islands of Tenerife also exhibit a topographic complexity and variety of micro-climates that might have favoured their role as micro-refugia against climate- or human-induced extinction (Harter et al. 2015).

In reviewing the role of Tenerife paleo-islands as refugia across several plant lineages, Trusty et al. (2005) found that species endemic to the paleo-islands often occupied a derived position in the lineage's phylogeny. This position argued against the idea of these massifs as
ancient refugia. However other studies, especially in animals, have reported ages for divergence events between taxa endemic to these paleo-islands, either at the species or at the infra-species level (Juan et al. 1996, 2000; Dimitrov et al. 2008; Macías-Hernández et al. 2013; Puppo et al. 2014;) that are contemporaneous or predate the age of merging of the precursor paleo-islands 3.5 million years ago (Ancochea et al. 1990). Few plant studies (Gómez et al. 2003; García-Verdugo et al. 2010) have focused on patterns of within-island genetic variation for widespread Canarian endemics, and none of them have provided estimates of lineage divergence times, which is necessary to relate within-island patterns to the island geological history. Species that are present in multiple islands ("multiple island endemics" or MIEs) are especially relevant to understand the role of paleo-islands as undisturbed areas that have acted as reservoirs and sources of genetic diversity not only within- but between-islands.

Here, we study patterns of genetic diversity and the demographic and spatial history of a multiple island endemic, Canarina canariensis (L.) Vatke. This "flagship" species of the Canary Islands, elected as its “national flower” (Kunkel 1991), is a diploid (2n = 34) herbaceous plant that grows mostly in cleared areas surrounding the endemic laurisilva forest. It presently occurs in the central and western Canary Islands: Gran Canaria, Tenerife, La Gomera, La Palma and El Hierro. C. canariensis is a herbaceous plant that occasionally climbs on nearby plants; it is pollinated by generalist birds (Rodríguez-Rodríguez & Valido 2011) and its fleshy fruits are dispersed by vertebrates (Valido et al. 2003, Rodríguez et al. 2008). Genus Canarina belongs to tribe Platycodoneae, a basal group within family Campanulaceae (Olesen et al. 2012; Mansion et al. 2012; Wang et al. 2013; Mairal et al. 2015). In addition to C. canariensis, the genus comprises two other species inhabiting the Afromontane forests of Eastern Africa, C. eminii and C. abyssinica, being this an example of a wide, continental-scale disjunction of 7000 km spanning across the Sahara. Mairal et al.
(2015) recently reconstructed the phylogeny and spatiotemporal evolution of *Canarina*. They inferred that *C. canariensis* diverged from its African relatives at the end of the Miocene (c. 7 Ma). This extraordinary temporal and spatial disjunction was explained as the result of vicariance and climate-driven extinction resulting in the fragmentation of an ancient widespread distribution. The colonization of the Canary Islands by the ancestors of *C. canariensis* apparently occurred much later, in the Pleistocene, probably from a now extinct and geographically closer North African population: the earliest event of population divergence is dated at only around 1 Ma (Late Pleistocene) and involved an east-west vicariance within Tenerife (Mairal *et al.* 2015).

This age and the presence of *Canarina canariensis* in several islands makes it an ideal candidate to evaluate patterns of within-island diversification in relation to the recent geological history of the archipelago. Our main aims were to: i) determine the geographic distribution of genetic variation within *Canarina canariensis*, ii) find evidence of extinction and diversification processes that may be related to geological events, iii) find ancestral areas and reconstruct inter-island migration events, iv) examine the putative role of the paleo-islands of Tenerife as refugia of genetic diversity, both relictual and recent.

Haplotype networks are commonly used in population-level studies because they provide a clearer picture of the reticulate relationships between genetic pools than a branching tree, especially when gene flow is present. These networks are often inferred using Statistical Parsimony (SP, Templeton *et al.* 1992) implemented in the software TCS (Clement *et al.* 2000), which allows estimation of the haplotype network while minimising the number of mutation events differing among haplotypes. However, this method fails to incorporate the uncertainty associated to the network inference, and therefore does not allow for statistical evaluation of alternative phylogeographic scenarios (Bloomquist *et al.* 2010). Moreover, unobserved events such as local population growth or past extinction of haplotypes may

This article is protected by copyright. All rights reserved.
This article is protected by copyright. All rights reserved.
followed those of Mairal et al. 2015. The sources of the material examined, the GenBank accession numbers and full references are detailed in Table S2.

**Haplotype analyses**

Sequences for each region were aligned using MAFFT 6.814b (Katoh et al. 2002), implemented in the software Geneious Pro 5.4.4. (Biomatters Ltd., Auckland, New Zealand). Sequences were checked and manually adjusted where necessary by following alignment rules described in Kelchner (2000). We analysed the three sequenced regions as three data partitions to perform phylogenetic analyses. MrModeltest v.2.2 software (Nylander, 2004) was used to determine the best fitting model of sequence evolution of each data partition.

Summary statistics for within-population genetic diversity were calculated as: the number of haplotypes $H(n)$, haplotype diversity ($H_d$), nucleotide diversity $\pi$, nucleotide heterozygosity $\theta$, $G_{ST}$ and the number of migrants per generation ($N_m$) were estimated for each population using DnaSP (version 5.10; Librado & Rozas 2009).

The relationships amongst lineages were investigated through haplotype network analysis, by using 6-12 individuals from different populations and examining the three sequenced regions. Genealogical relationships among haplotypes were inferred via the statistical parsimony algorithm (Templeton et al. 1992) implemented in TCS 1.21 (Clement et al. 2000). The number of mutational steps resulting from single substitutions among haplotypes was calculated with 95% confidence limits, gaps were represented as missing data.

The BPEC method (Manolopoulou et al. 2011; Manolopoulou & Emerson 2012) was implemented in the R package (Manolopoulou & Hille, in prep.) to identify genetically distinct geographical population clusters and ancestral locations. Like TCS, BPEC relies on parsimony in order to reduce the number of candidate trees to a manageable set. The BPEC
method, unlike Standard Parsimony, fits a prior over all possible trees in order to identify trees with high posterior probability in a fully model-based framework, thus accommodating for uncertainty in haplotype relationships, which is one of the main criticisms of TCS (Knowles 2008). Each possible tree defines a set of possible migration events that may have led to the observed population substructure. Different scenarios of trees and migration events are explored through Markov chain Monte Carlo (MCMC), similar to the method proposed by Sanmartín et al. (2008) for estimating rates of inter-island dispersal. Migration events were assumed to occur when a haplotype (with or without a mutation from its parent haplotype) migrates to a new geographical cluster. MCMC simultaneously estimates high probability trees, number of migration events and corresponding clusters. The method assumes that the migration rate and the population growth are constant. BPEC requires two main user-defined inputs: the maximum number of migrations (denoted as “MaxMig” in the software) and the parsimony relaxation parameter used to reconstruct the set of possible trees (denoted as “ds” in the software). MaxMig allows the user to set the upper bound for the number of migration events and hence the maximum number of clusters (MaxMig + 1). Larger values include more models but require much greater computation time. As the ds value is increased, the parsimony assumption is relaxed: if two observed sets of sequences have an unobserved intermediate missing sequence (an unobserved mutation), then any pair of sequences of distance \(1,\ldots, ds\) nucleotides will be considered as the “missing path”. Two MCMC chains were run for three million iterations. The results were stable, with ds > 3 having no effect on the inferred tree, and similarly, any number of migrations above 4 converging to a 5–cluster model. The phylogeographic clustering obtained was superimposed upon a haplotype tree, and used to estimate ancestral locations for migration events. As a further exploration, we divided the dataset into two groups (eastern and western clades – for details see coalescent dating section in results). Haplotypes sampled from Roque del Conde
were quite divergent; thus, separate analyses were run in which we included and excluded this location from the two groups. We ran BPEC analysis on each of these four datasets (eastern and western groups with and without Roque del Conde).

Haplotype divergence times were estimated in BEAST v.1.7 (Drummond and Rambaut 2007). We carried out a first analysis under a strict clock model and a coalescent constant population tree prior, using a secondary age estimate (Mairal et al. 2015; normal prior: mean = 0.76 Ma, standard deviation (SD) = 0.327 Ma) to calibrate the root node of the “C. canariensis dataset”; this included all haplotypes detected in our sampling (N = 10). This analysis gave us very large 95%HPD (High Posterior Density) credibility intervals and poor ESS for posterior age estimates, probably due to the low information content at the population level (see Results) and presence of a single calibration point. We carried out a second analysis applying the “nested dating approach” described in Mairal et al. (2015), in which a higher-level dataset including representatives of all three species of Canarina and 9 outgroup taxa was used to inform the clock rate of a linked population-level dataset (C. canariensis) under a mixed Yule-coalescent model (Ho et al. 2005; Pokorny et al. 2011). The higher-level dataset was calibrated with fossil-derived secondary age estimates (see Mairal et al. 2015), while the tree prior was unlinked to apply a coalescent constant size model to the population-level dataset and a stochastic birth-death (Yule) prior to the species-level one (Mairal et al. 2015). The clock model was set to an uncorrelated log-normal prior to accommodate the change in mutation rate from species to populations, with a uniform distribution for the ucld.mean ($10^{-4}$-$10^{-3}$) and a default exponential distribution for the ucld.stdev; the substitution model was set to GTR+G; choice of priors was based on Bayes Factor comparisons using the path sampling method in BEAST (Baele et al. 2012); see Table S3 to see results from exploratory analyses to assess the reliability of our date estimates with reference to these settings. Two MCMC chains were run for 50 million generations, sampling...
parameters every 1000 generation. We used Tracer v1.6 (Rambaut et al. 2013) to verify: whether a stationary distribution was attained; if there was convergence amongst chains; and that effective sample sizes (ESS values) were > 200 for all parameters. A 10% burn-in of the sampled populations was discarded (5 million). Post-burn-in trees were summarised into a maximum clade credibility tree using TreeAnnotator v.1.6.1, with mean values and 95% credible intervals for nodal ages, and were visualised in FigTree 1.3.1 (Rambaut and Drummond 2009). The resulting age estimates from this second analysis exhibited considerably larger ESS values and narrower 95%HPD intervals than in the first and are the ones reported here.

Demographic history.

Statistics used to describe demographic patterns may be biased by a strong genetic structure or lack of panmixia among populations. Since we detected strong genetic structure in our dataset (see results), we performed demographic analyses in subsetted datasets that were less genetically structured (including approximate panmitic populations). Three groups previously recognised by the haplotype network and BEAST analyses were used: ET-GC (including populations in east Tenerife and Gran Canaria), TENO-GO (including populations in Teno and La Gomera) and LP-EH (including populations in La Palma and El Hierro). For the same reasons as described above, the Roque del Conde population was included and discluded from the ET-GC and TENO-GO groups. Overall we ran the analyses on five groups.

We used three different approaches to infer the demographic processes shaping the genetic structure of *C. canariensis*. First of all, to test for evidence of population expansion, we carried out a neutrality test – Fu and Li´s tests (Fu & Li, 1993; Fu 1996) and Tajima´s $D$ test (Tajima, 1989) – for each population group. We used the DNAsp program,
version 5.0 (Librado & Rozas 2009) and assessed the significance of each test with 10,000 coalescent simulations. Secondly, we plotted the mismatch distribution for each group using the observed number of differences between all pairs of sequences with the ARLEQUIN v.3.0 software (Excoffier et al. 2005). The goodness-of-fit of the observed mismatch distribution to the theoretical distribution under a constant population size model was tested with the raggedness index ($HRag$) (Harpending et al. (1994)). Thirdly, i) we created the Extended Bayesian Skyline Plot (EBSP), implemented in BEAST, for each population group, ii) we performed EBSP analyses for each of the 16 population groups. For each group from i) and ii): two independent chains were run simultaneously for 150 million generations, sampling every 1000 generations; a strict clock model was used, whereas all other parameters were set identical to those described above for the nested dating analysis; the root node was calibrated using a normal prior with a mean age estimate and 95% High Posterior Density (HPD) credible intervals obtained from this analysis.

**AFLP fingerprinting**

For the AFLP analysis we used a total of 97 individuals from 10 populations, which covered all of the islands: one population from Gran Canaria, five from Tenerife, one from La Gomera, two from La Palma and one from El Hierro. Laboratory molecular protocols for the AFLP analysis (Vos et al. 1995) were implemented using the AFLP plant mapping kit (Applied Biosystems®). To select the appropriate primers, we first carried out a pilot study combining fluorophores and restriction enzymes for five geographically distant individuals (one per island), using in total 32 primer combinations. One sample from each individual was duplicated as a blind sample to test for reproducibility and contamination. Reproducibility and the number of alleles per sample were calculated by choosing three combinations of primers: 1-$EcoRI_{FAM}^{6}$-ACT/$MseI$-CAA, 2-$EcoRI_{VIC}^{AGG}$/MseI-CTA, and 3-$EcoRI_{VIC}^{AGG}$-ACT/$MseI$-CAA.
AGG/MseI-CTT and using the Genemapper v3.7 software (Applied Biosystems, Foster City, CA, USA). These three primer pairs showed high reproducibility and homogeneously scattered bands, and produced polymorphic AFLP profiles and clear fragments. In the digestion phase, samples of DNA were digested with the restriction enzymes EcoRI and MseI and linked to the primers EcoRI 5’-CTCGTAGACTGCGTACC-3’/5’-AATTGGTACGCAGTCTAC-3’) and MseI (5’- GACGATGAGTCCTGAC-3’/5’-ATCTCAGGACTCAT-3’). The three different AFLP reactions were: i) restriction and ligation in a single reaction; (ii) and (iii) consecutive PCR amplifications (preselective and selective). PCR products were checked on 1% agarose gels.

**AFLP data analyses**

The resulting AFLP fragments were analysed using the GeneMapper 3.7 software. Peaks were recorded in 100 to 500 base pairs ranges. Shorter fragments were discarded because the majority of this size class have a high chance of being non-homologous fragments (Vekemans et al. 2002). For each primer combination, an automated size detection and peak binning was employed followed by manual editing of bins to exclude shoulder peaks and unreliable loci (variation between replicates). Peak height data were then exported and loaded into the R package AFLPScore version 1.4a (Whitlock et al. 2008) and the AFLP profiles were scored and the error rates were estimated. These rates were below the critical bound of 5% indicated in previous reports (Bonin et al. 2004) for each primer combination. Data reliability was assessed through comparison of duplicates, from one or two individuals per population. Data were converted into binary presence/absence scores for each locus. The resulting AFLP presence/absence matrix was analysed using a selection of different analyses. The AFLPSURV v.1.0 software (Vekemans et al. 2002) was used to estimate demographic statistics such as Nei’s gene diversity (Hj), pairwise differentiation among subpopulations.
(F_{ST}) and the percentage of polymorphic fragments per population (P) (Nei and Li, 1979; Lynch and Michigan, 1994). This was done under the assumptions of either the Hardy-Weinberg equilibrium or partial self-fertilisation, based on a previous study on the reproductive biology of C. canariensis (Rodríguez-Rodríguez and Valido 2011). A Bayesian method in AFLPSURV was also used to estimate allelic frequencies through employment of a non-uniform prior distribution (Zhivotovsky 1999). Ten thousand permutations were run to calculate the F_{ST} parameter from which genetic distances between individuals, populations and geographic groups were calculated. To locate genetic clustering of individuals within the AFLP dataset, a pairwise similarity matrix for all individuals using the Dice’s coefficient as similarity distance was constructed, and the resulting matrix subjected to a Principal Coordinates Analysis (PCA) implemented in Ntsys v.2.1 (Rohlf 1998). Next, genetic relationships among samples were visualised in SplitsTree v.4.10 (Huson and Bryant 2006) using Neighbour-Net analysis through the use of the split decomposition method. Finally, to quantify the amount of genetic differentiation attributable to geographic and population subdivision, a hierarchical analysis of molecular variance was performed using ARLEQUIN v.3.0 software (Excoffier et al. 2005). Exploratory analyses were performed considering, alternately, islands and paleo-islands as geographical units in order to investigate the distribution of genetic variance attributable to oceanic barriers.

Bayesian clustering methods implemented in STRUCTURE v.2.3 (Pritchard et al. 2000; Falush et al. 2007) were used to assess the genetic structure of populations. This model-based approach assumes that loci are in Hardy-Weinberg equilibrium and linkage equilibrium within populations. Analyses were performed under admixture conditions and correlated allele frequencies between groups. 500,000 MCMC generations (plus a burn-in of 100,000) were run for K values of 1-10, with 10 repetitions for each. The most likely K value was determined by the method from Evanno et al. (2005), which is implemented in
STRUCTURE HARVESTER (Earl 2012). We explored other values of K to detect further genetic substructure of populations, especially in Tenerife. To test the effect of the spatial distance on the genetic structure of the populations, correlations between genetic (measured as $F_{ST}$) and spatial distances between pairs of populations were determined using the Mantel permutation procedure implemented in NTSYS v. 2.1. The genetic distance matrix used was based on the presence/absence matrix; the geographic distance matrix was based on the absolute distances between the geographic coordinates for each collected population. In addition, to identify possible geographic locations acting as major genetic barriers among C. canariensis populations, we computed barriers on a Delaunay triangulation using Monmonier’s algorithm in BARRIER v.2.2 (Manni et al. 2004). The significance was examined by the mean of 1000 bootstrapped distance matrices obtained using AFLPsurv. Only barriers with support greater than 96% were considered as significant.

RESULTS

Haplotype network analysis and coalescent dating

The pet $^{1365}$–pet $^{738}$ region consisted of 937 sites, rpl32–trnL$^{UAG}$ of 654 sites, and trnS$^{GCU}$–trnG$^{UCC}$ of 658 sites. The final concatenated data matrix consisted of 144 sequences and 2249 nucleotide sites, of which 9 were polymorphic (Table S4). Among the 144 individuals sampled from the 16 populations, we observed 10 different plastid DNA haplotypes (H1 to H10, Figure 2a) with haplotype diversity of 0.6632. There were three haplotypes that dominated and were clearly geographically delimited: H1, H9 and H3. The most frequently observed haplotype was H1 (52.08%), which was dominant in GC and east Tenerife (ET). Haplotype H9 was restricted to West Tenerife (WT) and LG. Haplotype H3 was present in the western islands of LP and EH. Populations with the commonest and rarest haplotypes (frequencies ≤ 2%) were in the most rugged areas of Tenerife: the paleo-

This article is protected by copyright. All rights reserved.
islands of Anaga (An-W, An-C, An-E – exclusive haplotypes H2 and H5), Teno (Ten-A, Ten-Ep – exclusive allele H8 and allele H10 shared with LG) and Roque del Conde (Conde – exclusive H7). The small western islands presented one unique haplotype, each: LP (LP-Lt – H4) and EH (H6).

Summary statistics for within-population genetic diversity are shown in Table 1. Populations situated in the three paleo-islands (Anaga (An-W), Teno (Ten-A) and Roque del Conde (Conde)) showed the highest values for the number of haplotypes H(n), haplotype diversity H(d), nucleotide diversity π and nucleotide heterozygosity θ. The least diverse population areas were found in Gran Canaria and outside of the paleo-islands regions of Tenerife. G_{ST} and Nm values indicated high genetic cohesion between some islands: east Tenerife populations were highly connected with Gran Canaria populations; west Tenerife with La Gomera; and La Palma with El Hierro (Table S5). Interestingly, the highest G_{ST} values and lowest Nm were found between populations situated east and west of Tenerife, with the exception of Anaga An-W. This is the only population in east Tenerife that showed some genetic cohesion with populations in west Tenerife, in particular with the Teno and Adeje massifs.

The haplotype network constructed with TCS (Fig. 2b) shows a loop involving three dominant haplotypes, with haplotype H1 as the ancestral root haplotype according to coalescent criteria on haplotype frequency (Templeton et al. 1992). Dominant haplotypes at each side of Tenerife (East H1 and West H9) were separated by five mutational steps, with haplotype H10 as intermediate. Haplotype H3, present in the western islands, was separated by four mutational steps from H1 and three from H9.

BPEC results are summarised in Table 2. With mutational step limit equal to 10 and the maximum number of migrations equal to 8, we allowed for high gene flow, genetic divergence and numerous hidden mutations, which separate the haplotype clusters.
clustering with high posterior probabilities (pp = 1) showed the existence of five clusters (contour regions in Figure 3a). The total number of haplotypes was 10 plus one missing haplotype (Fig. 3b). Location of the ancestral populations for migration events and inference of the root node were consistent across exploratory analyses of the complete dataset. The root node was inferred as missing (H11 – Fig. 3b), and carried a high uncertainty. Haplotypes H1, H2, H8 also carried significant posterior mass probabilities (Table 2). The most likely source for ancestral migration events within- and between-islands were populations located on: Conde (pp = 0.13), An-W (pp = 0.10) and LP-Lt (pp = 0.086). The subsetting analysis, including Conde, always recovered this population as the ancestral area with the highest probability. Without Conde, other paleo-island populations were recovered as ancestral areas: Ten-A for the Western clade; and the three populations of Anaga for the Eastern clade, with An-W the most probable (Table 2, Figure S1).

The Bayesian chronogram of haplotypes showed a geographical pattern of divergence congruent with the groups detected by TCS and BPEC (Figure 4). The crown age or first divergence event among Canarina haplotypes was estimated to be 0.878 Ma (95% HPD = 0.452-1.365), corresponding to the Mid-Pleistocene period; it divided haplotypes into an eastern and western clade. The first divergence event in the eastern clade was estimated at 0.255 Ma (95% HPD = 0.041-0.633), while that in the western clade was dated older: 0.522 Ma (95% HPD = 0.161-0.967), separating La Palma and El Hierro from Teno and La Gomera. Mean age values and 95% HPD intervals for the BEAST analysis are shown in Table S6.
**Demographic analyses**

Negative values of Fu’s and Tajima's indicated a recent population expansion, although these values were only significant for the Fu’s test of the East group (Table S7). The frequencies of pairwise differences in the mismatch distribution analysis resulted in unimodal distributions that were consistent with an expansion model. When we include the population of Conde in the East group, we obtained a second small peak. The raggedness statistics derived from the mismatch distribution were not significant and thus failed to reject the null hypothesis of recent population expansion. EBSP indicated a constant population size (Figure S2). EBSP for each of the individual populations also obtained a constant population size (results not shown).

**AFLP polymorphism, genetic diversity and structure.**

These results are summarised in Table 1. The final dataset, after scoring, comprised 572 loci from 97 individuals and 10 populations. For AFLP analyses we removed populations from Gran Canaria and Roque del Conde, which had been extracted with a different (less rigorous) method (CTAB) to standardise the quality of extracted DNA. Hj and P were higher in the La Gomera (GO) population, followed by the Teno (Ten-A) population in west Tenerife. The largest number of private fragments was detected in Tenerife (7 fragments). In La Gomera and El Hierro islands a single fragment was detected. In Tenerife, five private fragments were detected in the Anaga massif (specifically in the An-W population – four fragments) and two in the Teno massif (Table 1). F_{ST} values with the nuclear data (Table S8) were consistent with the results obtained from the chloroplast markers (see above G_{ST} and Nm), and showed the same genetic cohesions between-islands and paleo-islands.
A PCA analysis differentiated three groups (Figure S4): i) east Tenerife populations; ii) west Tenerife populations; iii) the An-W population, which demonstrate an intermediate position between i) and ii). The split network analysis outlined the same divergences between the two well separated groups, including the West group An-W population (Figure 5b).

Hierarchical AMOVA analyses showed the largest proportion of genetic variation to be found among groups 4 and 5 (Table 3 and S9). By analysing each island separately, we observed that differentiation within-islands (24.27%) is greater than between-islands (3.62%). A further analysis considering islands and paleo-islands as separate units showed that differentiation within-islands and within-paleo-islands is lower (11.30%) than between-islands and between-paleo-islands (16.63%). According to the method of Evanno et al. (2005), STRUCTURE indicated that the most likely number of genetic clusters K = 2 (ΔK=460) represented the optimal number of Bayesian groups within C. canariensis, separating the east populations from the west, but also detected the presence of admixture in the intermediate populations. K levels K = 4 and K = 5 (the latter identified as K = 6, with 5 defined clusters plus one "ghost" cluster with no individuals assigned so it was ignored, see Guillot et al. 2005) revealed a more complex genetic structure in these intermediate populations, resulting in additional clusters (Fig. 5a, Figure S5). The STRUCTURE analysis using only Tenerife populations also delimited two groups (inset K = 2 in Fig.5a, Fig. S5b). Overall, the results revealed a strong inter-island genetic structure (Gran Canaria – Anaga massif and La Palma – El Hierro), admixture phenomena (among An-W, Teno populations, and La Gomera), and an even more complex substructure within the island of Tenerife. The STRUCTURE sublevels (K = 4 and K = 5) were consistent with the AMOVA analyses. Only La Gomera differed between analyses; this is explained due to its high genetic variability and mixed composition. These groups are also congruent with the BARRIER

This article is protected by copyright. All rights reserved.
results, which revealed two or four major boundaries (Fig. 5a for K=2 and K = 5). All barriers showed values of 100% except LP-EH (96.8%). No linear relationship was found between pairwise $F_{ST}$ and geographical distance with the Mantel analysis ($r = 0.42, p = 1$).

**DISCUSSION**

*Paleo-islands as refugia against volcanic (catastrophic) extinctions*

The strong genetic structure often detected in island endemics has been postulated to respond to several interconnected physical and biotic factors (Stuessy *et al*. 2014); foremost among them are geographical isolation and extinction-recolonisation processes (Carson *et al*. 1990; Gübitz *et al*. 2005; Macías-Hernández *et al*. 2013). The strongest genetic structure and levels of genetic variation in *C. canariensis* were detected between two of the ancient paleo-islands that currently form Tenerife: Teno in the west and Anaga in the east. Populations in other islands were genetically associated to these two lineages: La Palma, La Gomera and El Hierro to Teno; and Gran Canaria to Anaga. The fact that these patterns are shared between chloroplast (cpDNA, Fig. 2) and nuclear (AFLPs, Fig. 5) markers supports a long history of isolation among populations (Zink & Barrowclough 2008). Furthermore, the hierarchical AMOVA analysis based on the AFLP data showed that genetic variation was notably higher among populations within Tenerife than those located in different islands (Table 3). This pattern has been found in a few other Canarian endemics (*Olea europaea guanchica* and *Pinus canariensis*), but usually not as marked as here (< 10%: Gómez *et al*. 2003; García-Verdugo *et al*. 2010; see review in García-Verdugo *et al*. 2014). Interestingly, considering the paleo-islands as separate areas reversed the pattern (Table 3). Together with our BARRIER analyses (Fig. 5a in K = 5), these results support the hypothesis that the geographical distribution of genetic diversity in *C. canariensis* is structured around the paleo-
islands, and that these ancient massifs could be considered as separate units in phylogeographic analyses, a sort of "islands within-islands". It indicates that geological barriers within islands – for example, those associated to catastrophic volcanic events – have probably been more important in structuring patterns of genetic differentiation within species than the oceanic barriers separating the islands (García-Verdugo et al. 2014).

Unfortunately, we could not obtain AFLP data for the population in Roque del Conde. Given the mixed composition of the chloroplast compartment in this population (Fig. 2), it is possible that inclusion of this population in our analysis would have led to higher levels of genetic admixture and lower K values in the STRUCTURE analyses – though given the marked east/west split among the remaining populations (Fig. 5), this is unlikely. On the other hand, the two most frequent haplotypes, H1 and H9, in the Conde population are also the most frequent or dominant within the eastern and western clades, respectively, while the divergence of these two haplotypes in the BEAST tree (Fig. 4) can be traced back to the basalmost split in *C. canariensis* (0.8 Ma). This, together with the presence of a unique haplotype (H7) and the fact that the Conde population is identified as the ancestral source of westward and eastward migration events in the BPEC analyses, suggests that the admixture detected in the chloroplast compartment for Roque del Conde is of ancient rather than recent origin. It is thus possible that including this population in the AFLP analysis would have increased rather than decreased levels of genetic diversity within Tenerife for *C. canariensis*.

Which might have caused this level of within-island genetic structure? Many Canarian plant and animal taxa include sister lineages endemic to the Tenerife paleo-islands (Juan et al. 2000; Báez et al. 2001), especially among laurel forest species from Teno and
Anaga (Table 4). While in some cases, the divergence between these taxa is found at the species level (e.g., *Trechus*, Contreras-Díaz et al. 2007; *Pericallis*, Jones et al. 2014) and/or predates the origin of the paleo-islands (*Micromeria*, Puppo et al. 2014; *Pholcus*, Dimitrov et al. 2008); in others, it is observed within species (e.g., *Tarentola delalandii*, Gübitz et al. 2000), and/or postdates the merging of the paleo-islands (*Eutrichopus*, Moya et al. 2004; *Calathus abaxoides*, Emerson et al. 1999). The fact that the pattern of divergence is not contemporary across taxa suggests that the role of Tenerife paleo-islands in structuring genetic variation has been long lasting. In *Canarina canariensis* the basal divergence between the basal eastern and western lineages was dated as 0.878 Ma (Fig. 4) substantially, postdating the age of the merging of the paleo-islands (c. 3.5 Ma, Ancochea et al. 1990). This ‘young’ east-west divergence might be explained by historical dispersal events between forest patches followed by in-situ diversification. Alternatively, it could be explained by allopatric speciation (vicariance) driven by catastrophic volcanic events within a previously more widespread distribution with low or no reconnection. This might explain the ‘young’ (< 3.5 Ma) east-west Tenerife divergence found in *Canarina*, and seen in other taxa (Table 4).

The reciprocal monophyly, similar haplotype diversity levels and deep temporal divergence found between the eastern and western clusters of *C. canariensis* in Tenerife (Fig. 4), favour the vicariant, allopatric scenario. In the last one million years, several major landslides and volcanic events have affected the north of Tenerife, extending from the summit to the coast (Boulesteix et al. 2012, 2013; Carracedo 2014), e.g. the Güimar and La Orotava megaslides (Fig. 1b). These events could have fragmented the ancient laurel forest corridor that extended across the northern flank of the island (Moya et al. 2004), interrupting gene flow within species associated to this laurel forest (e.g., *C. canariensis*) and promoting differentiation among populations. Several studies have supported debris avalanches as important factors driving diversification within terrestrial Canarian organisms (Juan et al.
2000; Brown et al. 2006; Macías-Hernández et al. 2013); others have reported a temporal divergence within northern Tenerife (Thorpe et al. 1996; Moya et al. 2004) that is similar to the one found in C. canariensis and corresponds well with the time of the La Orotava megaslides (Ancochea et al. 1990; Boulesteix et al. 2013).

Whereas central Tenerife was almost completely covered by catastrophic events until as recently as 0.13 Ma (Ancochea et al. 1999), the three paleo-islands of Teno, Anaga, and Roque del Conde, remained geologically stable since the mid-Pliocene (see Fig. 1). This suggests that these areas could have acted as refugia, allowing the survival of populations that otherwise disappeared in other parts of the island that were affected by volcanic activity. Support for this suggestion comes from the population genetic analysis. According to the central-marginal hypothesis (Eckert et al. 2008), spatial structure and genetic diversity should be higher in areas that have acted as refugia for the preservation of genotypes that went extinct in other areas and generally for the long-term persistence of populations (Hewitt 2000; Tzedakis et al. 2013; Gavin et al. 2014; Feliner 2014). The higher number of ancestral and endemic alleles, private fragments, and larger heterozygosity levels exhibited by the populations of the paleo-islands of Tenerife (Table 1), are congruent with the idea that these massifs acted as reservoirs of ancient genetic diversity and as refugia against volcanically induced extinction. Interestingly, La Gomera, an island that has been geologically quiescent since the Pliocene (Carracedo & Day, 2002), shows the highest Hj and percentage of polymorphic nuclear DNA fragments (Table 1), suggesting that this island might have acted in a similar way to the paleo-islands of Tenerife.

Further support for the extinction hypothesis comes from the BPEC analysis. Theoretical predictions of coalescent theory states that high frequency haplotypes have been present for a long time, and more recent ones are rare and derived from the commonest haplotypes (Posada & Krandall 2001). Additionally, a root haplotype is expected to have a
higher number of haplotype connections in the network, rather than being close to the tips. However, past extinction of haplotypes can obscure the inference, with younger haplotypes becoming the most prevalent, so accurately identifying the root haplotype is a challenging task. The fact that BPEC does not provide a single estimate of the haplotype network like Statistical Parsimony, but a finite (probability) distribution of haplotype trees - as well as the existence of the underlying migration model - allows this method to incorporate the uncertainty in the haplotype rooting. BPEC estimates the missing, extinct haplotype H11 as the most probable root of the haplotype network (Fig. 3b), although haplotypes H1, H2 and H8 are also associated with high posterior probabilities.

*Have paleo-islands acted as sources of genetic diversity within and towards other islands?*

The theory of Pleistocene climate refugia (Hewitt 2000) states that historically environmentally stable areas can act as sources of genetic diversity exporting migrant alleles to other, disturbed regions (Gavin et al. 2014). Paleo-islands could have played the same role in volcanic archipelagos, although in this case catastrophic geological events rather than climatic changes might be responsible for the observed patterns. BPEC provides support to this hypothesis, identifying the populations in the paleo-islands of Tenerife as the source areas of ancestral migration events to other adjacent islands, such as from Teno to La Gomera or from Anaga to Gran Canaria (Fig. 3a,b, Table 2). A third dispersal event from west Tenerife/La Gomera to La Palma is supported by the BEAST tree (Fig. 4), and a fourth dispersal event from La Palma to El Hierro is inferred by the BPEC analysis (Fig. 3). Similar patterns with the central Canaries as centres of dispersal events within the archipelago have been described in other animal and plant studies (Francisco-Ortega et al. 2002; Gómez et al. 2003; Sanmartín et al. 2008; Puppo et al. 2015; see also Mairal et al., 2015). Moreover, a spatio-temporal pattern of colonization comparable to *Canarina*, showing Tenerife as the
center of dispersal events to adjacent islands in the last 1 Ma, can be found in the Canarian lineage of *Cistus* (0.33 (0.88-0.07) Ma; Guzmán & Vargas 2010), *Cheirolophus* (1 Ma; Vitales et al. 2014a), *Cistus monspeliensis* (0.93-0.20 Ma; Fernández-Mazuecos & Vargas 2011), and *Gallotia galloti* (0.8-0-9 Ma; Cox et al. 2010). This might be a consequence of the central geographic position of Tenerife within the archipelago – acting as a crossroad for dispersal events – but also of the concentration of plant genetic diversity in the paleo-islands.

The case of Gran Canaria is especially interesting. It shares the same haplotype with Tenerife (H1, Figs 2, 3). Low haplotype diversity (Hn and Hd in Table 1) could be explained by a recent colonisation after a catastrophic event. The island was subject to intense volcanic activity during the Holocene (24 eruption events; Rodríguez-González et al. 2009), so extinction might explain its present low genetic diversity. Another possibility is related to the topography of Gran Canaria, where a network of ravines (locally known as "barrancos") connecting at their summits, could have facilitated gene flow among populations. Dispersal of *Canarina canariensis* seeds by *Gallotia* lizards (Rodríguez et al. 2008), probably using forest gaps and edges of roads as dispersal corridors (Delgado et al. 2007), might have helped to connect populations in the highly altered laurel forest of Gran Canaria.

In addition to dispersal events between-islands, the paleo-islands of Tenerife might have acted as sources of genetic diversity within Tenerife, exporting migrant alleles to other geologically unstable, disturbed areas. Our demographic analyses indicated a recent population expansion in two populations of east Tenerife close to Anaga (TF-Br and TF-Bj). Although this result should be taken with caution (the EBSP analysis supported a constant effective population size; Fig. S2), these two populations exhibited also a single cpDNA haplotype, which agrees with the idea of a recent colonisation. The areas where these populations are located (La Orotava Valley and Güimar Valley, respectively) have been
subjected to catastrophic volcanic events (Fig. 1b). Some authors (Thorpe et al. 1996; Gübitz et al. 2000) have proposed the existence of a corridor of suitable habitat along the northern coastal fringe of Tenerife to explain migrations of the reptiles Gallotia galloti and Tarentola delalandii from Anaga to the west. These dispersal events could also explain the patterns found in our BPEC analyses, which suggest Anaga populations as sources of migration events to other populations in eastern Tenerife and Gran Canaria (Fig. 3, Table 2).

Ongoing genetic connectivity between the populations in the paleo-islands of Tenerife and those from other areas is supported by the nuclear genome, which shows genetic admixture between west-Anaga, Teno and La Gomera (Fig. 5a). Admixture could be explained by the carrying of pollen by nectar-feeding birds between forest patches (e.g. ringing techniques have confirmed migration of the main pollinator (Phylloscopus canariensis) between Teno and Anaga; Alejandro González, pers. comm.). The fact that this connectivity is to some extent lost in the cpDNA might be explained by the cpDNA not being transported via pollen or, alternatively, by the small size and haploid nature of the chloroplast genome, which imply shorter coalescent times and less time to fix novel mutations for chloroplast markers (Avise 2000; Jakob & Blattner 2006). On the other hand, the widespread distribution of some cpDNA haplotypes across the archipelago (H1, H3 and H9, Fig. 2a) supports some gene flow among populations driven by seed dispersal. In the Canary Islands, birds have been cited as important vectors for the dispersal of fleshy fruits (Arevalo et al. 2007; Padilla et al. 2012), and the latter has been associated with frequent gene flow preventing speciation within widespread island endemics (García-Verdugo et al. 2014), such as in Canarina. This fits with what is known on the reproductive biology of C. canariensis, which is pollinated by generalist birds (Rodríguez-Rodríguez & Valido 2011), while its fleshy fruits are dispersed by vertebrates such as Gallotia lizards (Valido et al. 2003, Rodríguez et al. 2008). An additional factor to explain frequent dispersal between-islands are
the eustatic sea-level shifts during the Pleistocene, which might have decreased geographic distance between-islands (Rijsdijk et al. 2014).

Finally, in addition to exporting migrant alleles to other islands and disturbed areas, the paleo-islands of Tenerife might have acted as cradles or sources or new genetic diversity. The higher haplotype and nucleotide diversity and higher Nei's gene diversity Hj exhibited by populations in these areas (Table 1) agree with their role as ancient refugia but also as sources of novel genetic diversity. Maximum topographic complexity is one of the main factors explaining species richness and high speciation rates within-islands (Whittaker et al. 2007). The rugged nature of the Tenerife paleo-islands has likely promoted genetic differentiation within these massifs. For example, divergences found between populations in east and west forest ranges within Anaga have been explained by the existence of deep ravines and the volcanic arc of Taganana (Fig. 1b), acting as geographical barrier to gene flow (Macías-Hernández et al. 2013). These divergences were also detected in *C. canariensis*.

**CONCLUSIONS**

Traditionally, the distribution of genetic diversity within archipelagos is assumed to be structured around oceanic barriers, with between-island divergences expected to be larger than within-island differentiation. Here, we showed that within-island genetic patterns might be as strong as or stronger than those observed between-islands when they are associated with historical volcanic events. In *Canarina canariensis*, geographic patterns of genetic variation are structured around the paleo-islands of Tenerife, with a minor secondary effect due to oceanic barriers. Carine & Schaefer (2010) argued that although relatively short oceanic distances separate the Canary Islands, they might be responsible for the high diversity levels
found in the archipelago, acting as effective barriers to dispersal and promoting allopatric speciation. However, our results suggest that this hypothesis might not be valid for endemic species with widespread distributions across several islands (e.g., *Canarina*), for whom stretches of ocean are apparently less of a barrier than topographic relief within volcanic islands. Phylogeographic studies on other MIEs (multiple island endemics) are needed to confirm this hypothesis. The paleo-islands of Tenerife have probably acted as both genetic refugia and sources of new diversity within- and between-islands. The preservation of genotypes that became extinct everywhere else and the topographic complexity of the paleo-islands makes them potential “phylogeographical hotspots” (Médail & Diadema 2009) and reservoirs of unique genetic diversity, whose conservation should be prioritised.

**ACKNOWLEDGEMENTS**

We are grateful to Mitchell Cruzan and four anonymous reviewers, whose comments helped to significantly improve our manuscript. We would like to thank Alberto Herrero, Emilio Cano and Fatima Durán for their laboratory assistance. Field work could not have been conducted without the cooperation of the staff at the Jardín Botánico Canario Viera y Clavijo (Gran Canaria); Jacinto Leralta and Ángel Fernández (La Gomera); Félix Manuel Medina from the Cabildo of La Palma, and the Cabildo of Tenerife and El Hierro are thanked for helping with sampling permits and accommodation during field expeditions. We are grateful to Gonzalo Nieto-Feliner for a critical review of the manuscript. We also thank Águedo Marrero, Alfredo Valido, Aurelio Martín, Carlos García-Verdugo, Juli Caujapé-Castells, Lisa Pokorny, María Candelaria Rodríguez-Rodríguez, Marta Martínez, Moisés Soto, Nadir Álvarez, Nils Arrigo, Rubén Barone, Xavier Picó and Yulán Úbeda for all the fruitful discussions. This work was funded by Spanish Ministry of Science through projects CGL2009-1332-C03-01 and CGL2012-40129-C02-01). M. Alarcón was funded by the JAE-
Doc program (CSIC/FSE). MM and VC were supported by MINECO FPI predoctoral fellowships (BES-2010-037261 and BES-2013-065389, respectively).

REFERENCES


Manolopoulou I, Hille A (in prep.) BPEC: An R package for Bayesian Phylogeographic and Ecological Clustering.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.
DATA ACCESSIBILITY.
DNA sequences: Genbank accessions KP797991 to KP798432.
GenBank accessions, sampling locations and/or online-only appendices uploaded as online supplemental material.
NEXUS files for the concatenated cpDNA dataset and original AFLPs Matrix of *Canarina canariensis*, Dryad doi:10.5061/dryad.j4103.

AUTHOR CONTRIBUTIONS
M.M., M.A., J.J.A. and I.S. designed the study; M.M. carried out the field work, M.A. and J.J.A. contributed some samples; M.M. performed the research and analyzed the data under the supervision of M.A. and I.S.; I.M., V.C., I.S., and M.M. performed the BPEC analyses; M.M. wrote the paper with help from I.S, M.A and J.J.A. M.M. and I.S. revised the paper. All authors contributed with comments and approved the final version.

TABLES
Table 1. Descriptors of within-population genetic diversity in the cpDNA haplotypes and AFLP markers for each population studied of *C. canariensis*. Abbreviations: H(n): number of haplotypes; H(d): haplotype diversity; π: nucleotide diversity; θ: nucleotide heterozygosity; Hj (se): Nei's gene diversity (standard error). Geographical locations for population codes are shown in Figure 2.

<table>
<thead>
<tr>
<th>Island (Paleo-island)</th>
<th>Population</th>
<th>Nº samples</th>
<th>Haplotypes</th>
<th>H (n)</th>
<th>H (d)</th>
<th>π</th>
<th>θ</th>
<th>Nº samples</th>
<th>Nº of polymorphic fragments (% in brackets)</th>
<th>Hj (se)</th>
<th>Number of private fragments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gran Canaria</td>
<td>GC-Cc</td>
<td>8</td>
<td>H1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>409 (68.7)</td>
<td>0.2344 (0.0085)</td>
<td>0</td>
</tr>
<tr>
<td>Gran Canaria</td>
<td>GC-Tm</td>
<td>11</td>
<td>H1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Gran Canaria</td>
<td>GC-Es</td>
<td>10</td>
<td>H1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Gran Canaria</td>
<td>GC-Ag</td>
<td>11</td>
<td>H1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tenerife (Anaga)</td>
<td>An-W</td>
<td>6</td>
<td>H1, H2, H5, H9</td>
<td>4</td>
<td>0.8</td>
<td>0.00112</td>
<td>0.00136</td>
<td>7</td>
<td>401 (67.4)</td>
<td>0.2677 (0.0088)</td>
<td>4</td>
</tr>
<tr>
<td>Tenerife (Anaga)</td>
<td>An-E</td>
<td>10</td>
<td>H1, H2</td>
<td>2</td>
<td>0.2</td>
<td>0.00018</td>
<td>0.00031</td>
<td>10</td>
<td>345 (58.0)</td>
<td>0.2303 (0.0080)</td>
<td>0</td>
</tr>
<tr>
<td>Tenerife (Anaga)</td>
<td>An-C</td>
<td>11</td>
<td>H1, H5</td>
<td>2</td>
<td>0.327</td>
<td>0.00015</td>
<td>0.00015</td>
<td>11</td>
<td>375 (63.0)</td>
<td>0.2453 (0.0078)</td>
<td>1</td>
</tr>
<tr>
<td>Tenerife</td>
<td>TF-Bj</td>
<td>4</td>
<td>H1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tenerife</td>
<td>TF-Br</td>
<td>6</td>
<td>H1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tenerife (Roque del Conde)</td>
<td>Conde</td>
<td>11</td>
<td>H1, H9, H7</td>
<td>3</td>
<td>0.618</td>
<td>0.00121</td>
<td>0.00091</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Tenerife (Teno)</td>
<td>Ten-Ep</td>
<td>8</td>
<td>H9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>351 (59.0)</td>
<td>0.2449 (0.0082)</td>
<td>0</td>
</tr>
<tr>
<td>Tenerife (Teno)</td>
<td>Ten-A</td>
<td>7</td>
<td>H8, H9, H10</td>
<td>3</td>
<td>0.524</td>
<td>0.00038</td>
<td>0.00054</td>
<td>7</td>
<td>415 (69.7)</td>
<td>0.2997 (0.0088)</td>
<td>2</td>
</tr>
<tr>
<td>Gomera</td>
<td>GO</td>
<td>12</td>
<td>H8, H9</td>
<td>2</td>
<td>0.303</td>
<td>0.00013</td>
<td>0.00015</td>
<td>12</td>
<td>473 (79.5)</td>
<td>0.3012 (0.0077)</td>
<td>1</td>
</tr>
<tr>
<td>La Palma</td>
<td>LP-Ba</td>
<td>10</td>
<td>H3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>285 (47.9)</td>
<td>0.1923 (0.0080)</td>
<td>0</td>
</tr>
<tr>
<td>La Palma</td>
<td>LP-Lt</td>
<td>10</td>
<td>H3, H4</td>
<td>2</td>
<td>0.2</td>
<td>0.00009</td>
<td>0.00016</td>
<td>10</td>
<td>348 (58.5)</td>
<td>0.2309 (0.0079)</td>
<td>0</td>
</tr>
<tr>
<td>El Hierro</td>
<td>EH</td>
<td>9</td>
<td>H3, H6</td>
<td>2</td>
<td>0.5</td>
<td>0.00022</td>
<td>0.00016</td>
<td>9</td>
<td>303 (50.9)</td>
<td>0.2115 (0.0082)</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2. Summary of results from the Bayesian Phylogeographic and Ecological Clustering (BPEC) analyses. Selected values for parameters MaxMig (maximum number of migrations) and ds (relaxed parsimony assumption) are shown. Ancestral locations (Anc. Loc.) for migration events are shown with their associated posterior probabilities.

<table>
<thead>
<tr>
<th>Maxmig ds</th>
<th>Anc. Loc. 1</th>
<th>Anc. Loc. 2</th>
<th>Anc. Loc. 3</th>
<th>Root node probability</th>
<th>n° of clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>8</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western Clade (WT + LP + GO + EH)</td>
<td>4</td>
<td>3</td>
<td>Ten-A (0.40)</td>
<td>Lp-Lt (0.31)</td>
<td>EH (0.18)</td>
</tr>
<tr>
<td>Eastern Clade (ET + GC)</td>
<td>1</td>
<td>3</td>
<td>An-W (0.212)</td>
<td>An-C (0.122)</td>
<td>An-E (0.116)</td>
</tr>
<tr>
<td>Western Clade + Conde</td>
<td>4</td>
<td>3</td>
<td>Conde (0.416)</td>
<td>Ten-A (0.197)</td>
<td>GO (0.190)</td>
</tr>
<tr>
<td>Eastern Clade + Conde</td>
<td>1</td>
<td>3</td>
<td>Conde (0.23)</td>
<td>An-W (0.17)</td>
<td>An-C (0.086)</td>
</tr>
</tbody>
</table>

Table 3. Hierarchical analysis of molecular variance (AMOVA) for C. canariensis based on allelic variation at different levels: (A) among groups, (B) among populations within groups and (C) within populations. Geographical locations of populations are shown in Figure 2.

<table>
<thead>
<tr>
<th>AMOVA groups</th>
<th>Nº of groups (K)</th>
<th>Levels</th>
<th>F-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>[GC + Anaga] [An-W + Ten-A]</td>
<td>4</td>
<td>17.44</td>
<td>13.06</td>
</tr>
<tr>
<td>[GO + LP + EH]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[GC+ An-E] [An-C]</td>
<td>5</td>
<td>18.92</td>
<td>11.21</td>
</tr>
<tr>
<td>[Ten-Ep] [An-W + Ten-A] [GO + LP + EH]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[GC]</td>
<td>TF</td>
<td>GO</td>
<td>LP</td>
</tr>
<tr>
<td>[GC]</td>
<td>An-E + An-C</td>
<td>An-W</td>
<td>Ten-Ep</td>
</tr>
</tbody>
</table>
Table 4. Phylogeographic breaks with divergence times reported in the literature between the paleo-islands of Tenerife.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Phylogeographical disjunction and groups</th>
<th>Divergence (Ma)</th>
<th>Markers</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species complex in darkling beetle <em>Pimelia</em></td>
<td>East Tenerife / West Tenerife</td>
<td>5.5 Ma</td>
<td>Mitochondrial (COI) and nuclear (ITS-1)</td>
<td>Juan et al. 1996</td>
</tr>
<tr>
<td>Lizard <em>Gallotia galloti</em></td>
<td>Western/ North-eastern lineages</td>
<td>0.7 Ma</td>
<td>Mitochondrial cytochrome b</td>
<td>Thorpe et al. 1996</td>
</tr>
<tr>
<td>Beetle <em>Calathus abaxoides</em></td>
<td>Teno / Anaga</td>
<td>350.000 years</td>
<td>Two mitochondrial (COI and COII)</td>
<td>Emerson et al. 1999</td>
</tr>
<tr>
<td>Skink <em>Chalcides viridanus</em></td>
<td>Teno/ Anaga</td>
<td>1.1 Ma</td>
<td>Two mitochondrial (12S and 16S)</td>
<td>Brown et al. 2000</td>
</tr>
<tr>
<td>Gecko <em>Tarentola delalandii</em></td>
<td>Clade 1. Teno + Roque del Conde/Anaga . Clade 2. Teno/Roque del Conde</td>
<td>Clade 1. 9–10 Ma Clade 2. 7.6 Ma</td>
<td>Mitochondrial cytochrome b</td>
<td>Gübitz et al. 2000</td>
</tr>
<tr>
<td>Mite <em>Steganacarus carlosi</em></td>
<td>Clade 1. Teno/Anaga. Clade 2. Roque del Conde / Anaga</td>
<td>Clade 1. 3.2 Ma Clade 2. 25 - 3.6 Ma</td>
<td>Mitochondrial (12S and 16S)</td>
<td>Salomone et al. 2002</td>
</tr>
<tr>
<td>Beetle <em>Eutrichopus</em></td>
<td>Teno (E. gonzalezi) / Anaga (E. canariensis)</td>
<td>0.7 Ma</td>
<td>Mitochondrial COI marker</td>
<td>Moya et al. 2004</td>
</tr>
<tr>
<td>Beetle <em>Trechus</em></td>
<td>Clade 1. Anaga + Teno (T.antonii, T. tenoensis, T.felis) / Anaga (T. fortunatus) Clade 2. “T.flavocintus”; Teno / East Tenerife</td>
<td>Clade 1. 1.73 Ma (HPD:1.48–2.01) Clade 2. aprox. 0.75 Ma</td>
<td>Part of mitochondrial genes cytochrome oxidase I and II (Cox1 and Cox2),and nuclear (ITS 2)</td>
<td>Contreras-Díaz et al. 2007</td>
</tr>
<tr>
<td>Spider <em>Pholcus</em></td>
<td>Anaga (P. malpaisensis, P. knoeseli) / Teno (P. intricus, P. muscaenisa / P. tenerifensis and P. roquensis)</td>
<td>3.93 Ma (HPD: 2.2–5.88)</td>
<td>Four mitochondrial (CO1, 16S, NADH and tRNAleu), Morphological data.</td>
<td>Dimitrov et al. 2008</td>
</tr>
<tr>
<td>Grasshoper <em>Arminda brunneri</em></td>
<td>Anaga + Güimar / Teno + Roque del Conde</td>
<td>1 – 0.17 Ma</td>
<td>Two mitochondrial (12s rRNA, ND5) and two nuclear gene fragments (28s rRNA, ITS2)</td>
<td>Hochkirch &amp; Goerzig 2009</td>
</tr>
<tr>
<td>Spider <em>Dysdera verneaut</em></td>
<td>Teno / Anaga</td>
<td>3.94 Ma (HPD: 5.1-2.7)</td>
<td>mitochondrial (cox1)</td>
<td>Macías-Hernández et al. 2013</td>
</tr>
<tr>
<td>Plant <em>Pericallis</em></td>
<td>Teno (M. echinata) / Anaga (P. nusslaginis)</td>
<td>2.87 Ma (HPD: 1.55–4.76)</td>
<td>nuclear ITS</td>
<td>Jones et al. 2014</td>
</tr>
<tr>
<td>Plant <em>Micromeria</em></td>
<td>Teno (M. densiflora) / Anaga (M. tenerifae, M. glomerata and M. rivas-martinezii) + Central group</td>
<td>5.2 Ma</td>
<td>8 nuclear loci</td>
<td>Puppo et al. 2014</td>
</tr>
</tbody>
</table>

Figure Captions

**Figure 1.** a) Geographic map of the Canary Island archipelago, with the code for each island used in the text and their age of emergence (million years, Ma). b) Map of Tenerife showing main Pleistocene landslides and other geological events cited in the text. The old basaltic series corresponding to the paleo-islands are shown in the three contoured massifs at the edges of the island. Lines show valleys created after landslides. Dashed lines indicate the post-erosional Las Cañadas volcano. Arrows indicate the directions of the landslides. The graph on the right shows the main Tenerife catastrophic events that are commented in the text; the numbers in the legend refer to those in the map. Ages and maps adapted from: Carracedo et al. 1998; Ancochea et al. 1999; Cantagrel et al. 1999; Guillou et al. 2004; Boulesteix et al. 2012; Zaczek et al. 2015.

**Figure 2.** Haplotype distribution and network inferred for the chloroplast markers by TCS. a) Pie charts show the geographical location of populations and the frequency of occurrence of each haplotype, circle size is proportional to population size. Population codes are given in...
Table S2. b) Statistical Parsimony network inferred by TCS. Each haplotype is shown in a different colour; circle size is proportional to its frequency among populations.

**Figure 3.** Results from the Bayesian Phylogeographic and Ecological Clustering (BPEC) analyses for the cpDNA markers. a) Phylogeographical clusters (coloured blobs) and ancestral location for migration events (denoted by arrows). The contour regions are centred at the 'centre' of each population cluster, and the shaded areas show the radius of 50% concentration contours around it. Locations situated beyond the clusters could also belong to these clusters, but with low probability; in the case of Conde, it suggests a mixed composition, with this population as source of migrant alleles to the east and western clusters. b) Haplotype network receiving the highest posterior probability. The small black circle H11 indicates an unobserved (missing) ancestral haplotype.

**Figure 4.** Maximum Clade Credibility (MCC) tree obtained from the BEAST analysis of cpDNA haplotypes, showing mean ages (above branches) and 95% HPD credible intervals. Numbers below branches indicate Bayesian posterior clade support values. Codes for tips (H1 to H10) correspond to the haplotypes shown in Fig. 2.

**Figure 5.** Results from the analysis of AFLP markers. a) Histograms showing the Bayesian clustering of individuals within populations (STRUCTURE); colours represent the proportion of individual membership to each inferred Bayesian group. Dotted lines indicate barriers to gene flow and their percentage, as inferred by BARRIER. b) Split network with points coloured according to location, as indicated in the legend. Codes for populations found in Table S2.

**SUPPORTING INFORMATION**
Additional supporting information may be found in the online version of this article.
Figure 3

![a) and b) diagrams showing geological features and locations]

Figure 4

![Diagram showing evolutionary relationships and ages]

This article is protected by copyright. All rights reserved.
Figure 5